Section 4.0

Steam Metabolism

4.1 Introduction

Whole stream metabolism is a measurement of ecosystem function that includes ecosystem-scale rates of photosynthesis (gross primary production, GPP) and respiration (community respiration, CR). The relative rates of GPP and CR in an ecosystem identifies the basal source of energy supporting the aquatic food web: allochthonous (from outside the system) or autochthonous (produced within the system). Stream metabolism, and the calculation thereof, is based on the premise that changes in Dissolved oxygen (DO) concentrations—between daytime highs to nighttime lows—are the result of photosynthesis (biologic production of O_2), respiration (biologic consumption of O_2) and reaeration (physical exchange with the atmosphere), as given by the following equation:

$$\Delta DO = GPP - \boldsymbol{CR} \pm \boldsymbol{G}$$

Where:

ΔDO = Change in DO concentration

GPP = Gross Primary Production

CR = Community Respiration

G= Reaeration x DO Deficit

Stream metabolism has been used to investigate rates of GPP and CR since the pioneering work of Odum (1956). Since that time, metabolism estimation has become a more practical metric with the availability of high quality, relatively low cost DO sensors and data loggers. Using oxygen sensors, *in situ,* or "free-water" metabolism techniques, have a number of advantages over mesocosm experiments as one does not have to consider container effects. Using free-water techniques also avoids scaling issues and benthic substrate heterogeneity that occur with chamber investigations. Free-water metabolism estimates integrate all the metabolic processes and surface water-benthos interactions that occur over an entire stream reach (Young et al. 2008, Izagirre et al. 2008).

Aquatic ecologists have subsequently investigated both natural and anthropogenic controls on whole stream metabolism such as geography (Hill et al. 2000, Bernot et al. 2010) land use practices (Young and Huryn 1999, Houser et al. 2005) and riparian disturbance (McTammany et al. 2007). Others have investigated how stream metabolic rates influence ecological processes, such as nutrient processing (Hall and Tank 2003) and ecosystem structure (Sabater et al. 2002). From these relationships it has been suggested that whole stream metabolism is

potentially an excellent indicator of stream health because metabolism incorporates the interactions among numerous factors that influence the chemical, physical and biological integrity of streams, including geomorphology, hydrology, riparian vegetation, in-stream vegetation, climate, biology and chemistry of an entire stream reach (Mulholland et al. 2005, Grace and Imberger 2006, Young et al. 2008).

We measured whole stream metabolism in 49 stream reaches along a gradient of ambient nutrient concentrations to evaluate the potential use of stream metabolism as a functional indicator of nutrient enrichment. Accordingly, we compared daily rates of GPP and CR to nutrient concentrations and possible covariates (e.g., stream slope, shading and turbidity). Secondly, we compared daily rates of GPP and CR to DO criteria used by the State of Utah to assess if there were potentially deleterious impacts to stream biota associated with increased rates of GPP or CR. Lastly, we developed multivariate models to determine the most important physical covariates that influence GPP and CR and built a decision making matrix for water resource management.

4.2 Methods

We conducted whole stream metabolism estimates independently at two locations at each reference site, and at two locations above and below each Publicly Owned Treatment Works (POTW) site (Fig 2.1 Section 2). At each site we deployed a water quality probe (YSI 6600V2 or 600 OMS V2) to measure dissolved oxygen (DO) and temperature at five-minute intervals for a minimum of 48 hours. Solar radiation data were collected from the closest available weather station (mesowest.utah.edu). Surface water nutrients were collected at deployment and retrieval and were analyzed for total nitrogen (TN) and phosphorus (TP) at the Aquatic Biogeochemistry Laboratory at Utah State University (Valderamma 1981). We calculated stream metabolism using an open water method with reaeration (K) as a free parameter (Hall 2011, unpublished work) based on the following equation derived from Van de Bogert et al. (2007) (see Table 4.1 for symbol definitions):

$$O_t = O_{t-1} + \left(\frac{GPP\Delta t}{z} \times \frac{Light_t}{\sum Light} + \frac{CR\Delta t}{z} + K\,(O_{sat} - O_{t-1})\Delta t \right)$$

The model adjusts GPP and CR at each time step to fit the oxygen data using non-linear minimization (R function `nlm`) of the maximum likelihood accuracy estimates. In this equation, K can be modeled as a free parameter from the oxygen data simultaneously with GPP and CR. In rare cases where K could not be modeled accurately we had to constrain K with values calculated from nighttime regression (Grace & Imberger 2006) to improve model performance.

We used linear regression to evaluate the relationship between the nutrients (TN and TP) and the metabolic response rates GPP and CR. We then used a nonparametric deviance reduction (NDR, package `rpart`) procedure using least squares fitting to determine significant thresholds of TN and TP to separate GPP and CR into distinct groups based on TN and TP concentrations (Low, Medium and High). We then used ANOVAs followed by post-hoc Tukey's Honestly Significant Difference (HSD) to determine if there were significant differences ($p<0.05$) in daily rates of GPP and CR among the three groups. If daily rates of GPP and CR significantly differed among the three groups (Low, Medium and High) then the threshold values that define these groups provides nutrient concentrations that are generally associated with increased stream metabolism.

Table 4.2. Minimum and 30-day average dissolved oxygen standards listed by aquatic life beneficial use (UAC R317-2-14).

Aquatic Life	Minimum DO Standard	30-day Average DO Standard
Coldwater Fish	4.0	6.5
Warmwater Fish	3.0	5.5
Nongame Fish	3.0	5.0

In order to validate metabolism as a functional indicator of nutrient enrichment we evaluated whether metabolic rates were also associated with other measures of stream condition. We used the previously established GPP and CR nutrient enrichment groups to define three groups where GPP and CR significantly differed. We then used the same dissolved oxygen (DO) data used to calculate stream metabolism to calculate the percentage of time a site exceeded the daily minimum DO concentration observed at each location. To provide context, the DO data were compared against two of Utah's DO criteria—the daily minimum and minimum 30-day average—as estimates of acute and chronic effects to stream biota (UAC R317-2-14, www.rules.utah.gov/publicat/code.htm). Water quality violations were determined by assigning the appropriate DO standard to the most sensitive beneficial use for each water body (Table 4.2). We binned rates of GPP and CR into three groups based on the median values that were observed in each of the three nutrient groups. We then compared these measures of oxygen standards among the three GPP and CR groups (Good, Fair and Poor) using an ANOVA ($p<0.05$) and a *post-hoc* Tukey's HSD test. If the GPP and CR groups are significantly different, in terms of minimum DO and percent DO water quality criteria violations, then stream metabolism indicators are directly coupled to measures of stream health. Furthermore, any relationships between metabolism and standards violations is causal because metabolism metrics are calculated from changes in DO.

To compare the relative importance of in stream nutrients versus other potentially compounding stream characteristics on controlling daily rates

Table 4.1. Definitions of symbols in equation 2.

Symbol	Definition	Units
CR	Communtiy Respiration	g O_2/m^2/day
GPP	Gross Primary Productivity	g O_2/m^2/day
K	Reaeration coefficient	day^{-1}
Light	Solar radiation or PAR	parameter dependent
O	Dissolved oxygen	mg/l
O_{sat}	Oxygen saturation	mg/l
t	Time	fraction of day
z	Mean stream depth	meters

of GPP and CR, we used multivariate Random Forest analyses (Breiman 2001, R package `randomForest`). We compared available data for 20 physical factors that are known to—or have been suggested to— control whole stream metabolism obtained from GIS (USGS StreamStats), on-site physical habitat surveys (USEPA 2009), and water quality samples (Appendix A). We ran random forest regression on all variables and then selected the best performing variables (based on percent increase of mean square error) and re-ran the analyses to create the most parsimonious model possible. If the best subset of variables random forest model performed as well as the all variables random forest model (based on the pseudo-r^2 fitness statistic) then we considered the best subset model successful. The goal of this subset model was to find a few important variables that controlled GPP and CR that could be collected along with nutrients to increase confidence in decision making for impairments. All analyses were conducted in R v2.15.0 (R Core Development Team, 2012).

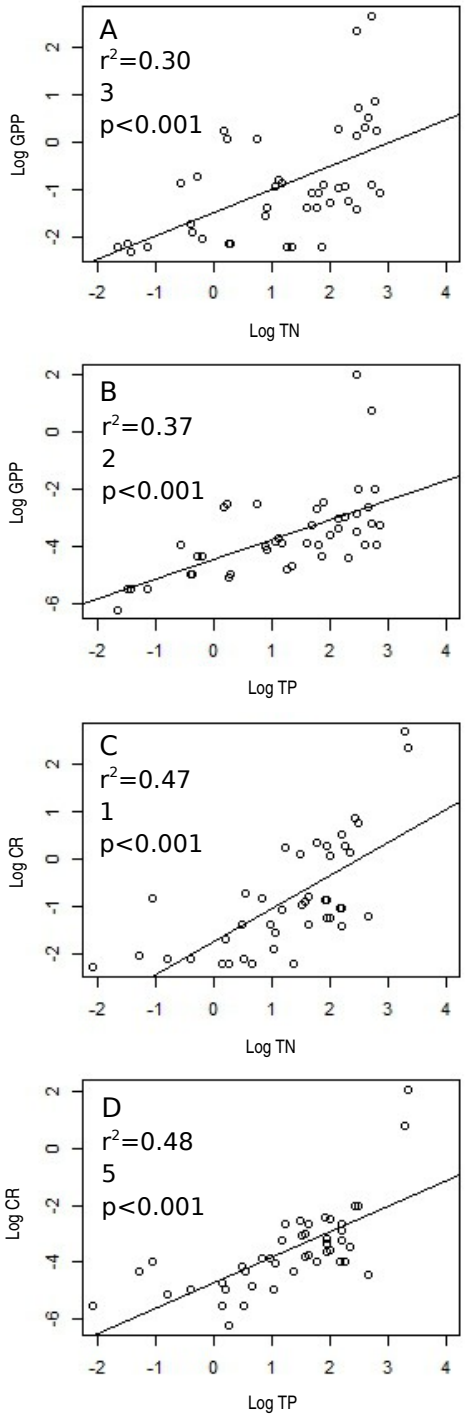

Fig 4.1. Linear Regression analysis between gross primary production (GPP) and total nitrogen (TN) (A) and total phosphorus (TP)(B) and between Community respiration (CR) and TN (C) and TP (D).

4.3 Results

Early exploratory analyses revealed metabolism rates were suppressed at highly turbid sites as were relationships between nutrients and rates of GPP and CR. Distributions of turbidity data revealed five highly turbid outliers with a turbidity of greater than 75 ntu. These outliers were excluded from all subsequent analyses. Nevertheless, a broad nutrient gradient (TN 0.10-14.37 mg/l and TP 0.002-7.65 mg/l) remained at the 44 stream reaches evaluated for all subsequent analyses. Future research will be required to determine if it is possible to develop defensible metabolism indicators for highly turbid streams.

We ran simple linear regression across all remaining sites (n=44) to determine the relationship between nutrients (TN and TP) and functional responses (GPP and CR (g $0_2/m^2/day$)). Across all sites GPP was positively related to both TN (r^2 = 0.303, p<0.001, Fig 4.1A) and TP (r^2 = 0.372, p<0.001 Fig 4.1B). CR

was more strongly related to nutrients than GPP for both TN (r^2 = 0.471, p<0.001, Fig 4.1C) and TP (r^2 = 0.485, p<0.001, Fig 4.1D).

Nutrient Thresholds. We used nonparametric deviance reduction (NDR, Qian et al. 2003) models to determine thresholds of TN and TP that best divide GPP and CR into relatively homogenous groups. The models identified three distinct groups with differing TN and TP concentrations (hereafter Low, Medium and High groups, Table 3). TN and TP nutrient groups generally corresponded with measures of stream metabolism. Among all sites GPP and CR rates differed among the three nutrient groups for both TN (ANOVA, p<0.001) and TP (ANOVA, p<0.001)(Fig 4.2). Mean GPP rates differed among both TN nutrient groups (Low TN=2.43±3.27 (standard deviation), Medium =6.57±4.9, High=13.19±2.59) as well as TP groups (TP-GPP rates Low=3.62±4.74, Medium =7.48±4.75, High=13.86±2.29) (Fig 4.2). CR daily rates also differed among nutrient groups established for TN (TN-CR rates Low=2.05±2.28, Medium =5.78±3.29, High=14.35±9.35) and TP concentrations (TP-CR Low= 3.13± 3.81, Medium =6.05±2.31, High=19.66±9.25) (Fig 4.2). The breakpoints in CR based on TN and TP that were often, but not always, identical to those for GPP. However, differences among GPP groups are probably being driven by the low nutrient group (Tukey's post hoc tests) for both TN and TP. Similarly CR differences seem to be a result of the small low nutrient group for TP, whereas each of the TN nutrient groups had distinctly different CR rates (Fig 4.2).

Table 4.3. Nutrient concentration thresholds used to groups sites to compare daily rates of GPP and CR and GPP and CR thresholds.

Nutrient	Nutrient Group Thresholds	Functional Indicator	Indicator Group Thresholds
TN (mg/l)	Low < 0.24 > Medium < 1.28 > High	GPP (g O_2/m^2/day)	Good < 6.0 > Fair < 10.0 > Poor
TP (mg/l)	Low < 0.02 > Medium < 0.09 > High	CR (g O_2/m^2/day)	Good < 5.0 > Fair < 9.0 > Poor

Metabolism Thresholds. Daily stream metabolic rates corresponded closely with the frequency of DO water quality criteria violations. We organized sites into three groups with differing GPP and CR rates (hereafter Good, Fair and Poor, Table 3) based on the differences in metabolic rates among nutrient groups. Overall, we found significant among-group differences in minimum daily DO concentration for GPP and CR (ANOVA, GPP p<0.001 and CR p<0.001). For GPP, we found significant differences in minimum DO concentrations between the GPP Good and Fair groups (Tukey's HSD p<0.001) and between the Good and Poor GPP groups (p<0.001), whereas there was no significant difference between the Fair and Poor GPP groups (p=0.06). We found the same pattern among the CR groups where Good and Fair (p<0.001) and Good and Poor (p<0.001) but no difference between Fair and Poor groups (p=0.98).

We found significant differences among GPP and CR groups and relative frequency of samples that exceeded minimum DO water quality criteria (ANOVA, GPP $p<0.001$ and ER $p=0.018$). For GPP, we found these exceedences differed between GPP Good and Fair (Tukey's HSD $p=0.001$) and the Good and Poor GPP groups ($p=0.02$), whereas no significant difference were observed between the Fair and Poor GPP groups ($p=0.06$). We found slightly different pattern among the CR groups where Good and Fair were also significantly different ($p=0.03$), but the Good and Poor ($p=0.12$) and the Fair and Poor groups ($p=0.83$) showed no significant differences. The Fair and Poor GPP and

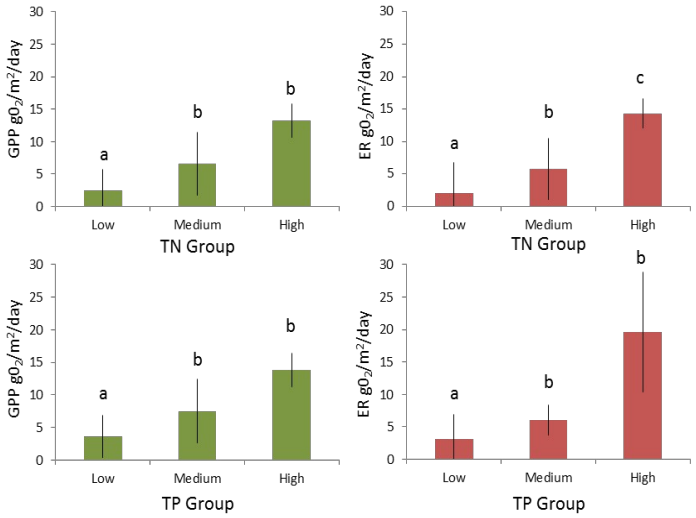

Fig.4.2. Bar chart comparing daily rates of GPP (green) and CR (red) among low, medium and high concentration sites for TN and TP. Thresholds for groups are shown in Table 2. Letters above bars indicate significant differences (Tukey's HSD $p<0.05$)

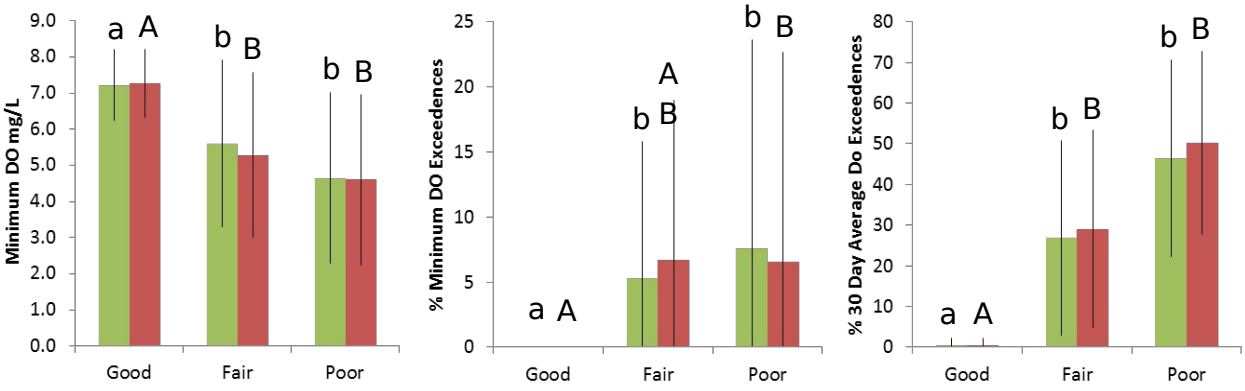

Fig 4.3. Comparisons of three measure of oxygen dynamics under different rates of GPP and CR(Good, Fair, Poor, Table 2). Lower case letters indicate significant differences of GPP groups and upper case indicates significant difference of CR groups determined by an ANOVA and *post-Hoc* Tukey's HSD. Error bars are standard deviation.

CR groups had a number of sites with exceedences, but also a large number with no exceedences, which lead to large within-group variation in minimum DO criteria violations (Fig 4.3).

In order to estimate the chronic effects of low DO we assessed the relationship among CR and GPP rates and the frequency of DO observations that fell below the 30-day average DO criterion assigned to the beneficial use of each site.

We found significant differences among GPP and CR groups and percentage of samples that exceeded the 30- day average minimum DO water quality standards (ANOVA, GPP $p<0.001$ and CR $p=0.018$). With GPP we found significant differences between Good and Fair groups (Tukey's HSD $p<0.001$) and the Good and Poor groups ($p<0.02$), whereas we found no significant difference between the Fair and Poor GPP groups ($p=0.26$). We found a similar relationship with the CR groups where Good and Fair and Good and Poor were significantly different ($p<0.001$ and $p<0.001$), but the Fair and Poor showed no significant difference ($p=0.25$, Fig 4.3).

Physical Covariates. We ran random forest regression models separately for GPP and CR with 20 variables (Appendix A) from water quality samples, GIS analyses and site specific habitat metrics. With all variables our model performed well in predicting GPP (mean squared residuals = 12.7, pseudo $r^2=0.54$) and CR (mean squared residuals=18.3, pseudo $r^2=0.45$, ntree=50,000). Four of the top five predictor variables were the same for GPP and CR as measured by increases of mean square error (MSE) resulting from randomly assigning values among observations. The top predictor variables for GPP were stream slope (MSE=103.9), stream shading (103.3), basin slope (74.9), TN (73.4) and TP (72.6). We found similar variables that were important for predicting CR including shading (MSE=70.2), TN (68.9), stream slope (63.2), mean stream depth (53.8) and TP (51.4). We ran random forest regression again with only the top four variables that were found in GPP and CR to compare overall model performance (stream slope, shading, TN and TP). We found that the model performed just as well with only the top four variables for GPP (mean squared residuals = 13.1, pseudo $r^2=0.53$) and CR (mean squared residuals = 16.2, pseudo $r^2=0.51$).

To verify the results of our random forest model we ran simple linear regression between major covariates and GPP and CR. There was a significant explanatory relationship between channel slope (%) and GPP (linear regression, $r^2=0.472$) and CR ($r^2=0.436$, data not shown). We ran NDR with GPP and CR and slope as the explanatory variable and we did find significant thresholds at ~1% slope. Channel shading measured as percent coverage (for methods see USEPA 2009, page 150, center densitometer readings) had a significant linear relationship with daily rates of GPP ($r^2=0.207$) and CR ($r^2=0.156$, data not shown). Using NDR we found a distinct threshold among rates of GPP and CR where shading equals ~ 11%. We found that streams with channel shading less than 11% had greater mean daily rates of GPP (9.3±5.6 to 3.99±4.1) and CR (8.10±5.5 to 4.31±4.1).

4.4 Discussion

A

Nutrient Thresholds. Using daily rates of GPP and CR we found two thresholds of TN and TP that can be used as indicators of where nutrient enrichment generally alters stream metabolic functions (Table 4.3). TN values of 0.24 mg/L and 1.28 mg/L and TP values of 0.02 mg/L and 0.09 mg/L separate low, medium, and high rates of both GPP and CR. These thresholds can be used by resource managers to evaluate where more intensive, follow-up sampling efforts are warranted. These values, along with other structural and functional indicators will

B

be used to create multiple lines of evidence of nutrient concentrations that can be developed into Statewide numeric nutrient criteria (See Section 8). We did not find many significant trends among common stream metabolism metrics such as production to respiration ratios (P:R) or net ecosystem metabolism (NEM, production-respiration). We did notice that as total phosphorus concentrations increased at a site NEM became more negative (data not shown). This same trend was not seen with total nitrogen concentrations suggesting an increasing heterotrophic response, relative to autotrophic response, to increasing phosphorus.

Metabolism Thresholds. One of the most direct and well known pathways between excess nutrients and deleterious effects on stream biota is through altering diel oxygen dynamics via increased autotrophic or heterotrophic productivity. Stream metabolism is an ideal metric to evaluate those effects because it directly quantifies these processes. Toward that end, by statistically binning daily rates of metabolism into three categories we were able to show a significant difference among the absolute minimum DO observed among sites and percent of times that DO observations were lower than minimum DO criteria (Fig 4.3).

Minimum DO standards are developed to protect aquatic life for each of the beneficial uses in the State. The three beneficial uses of streams used in this study were coldwater fisheries (3A), warmwater fisheries (3B) and non-game fish fisheries (3C) (UAC R317-2-6). Each of these beneficial uses has a different value for minimum DO concentrations based on the sensitivity of fishes, and other organisms in their food web, that are found in each type of fishery. Two sites were assigned with a beneficial use of severely habitat limited (3E) which does not have any DO standards and were given the least restrictive DO standards (3C beneficial use) for the purpose of these comparisons. The minimum DO water quality standard is the acute value where at any time the concentration in a stream falls below this value there would be a water quality violation. In our dataset only 11% of the sites sampled had a violation in the minimum DO standard criterion. However, these data may underestimate DO violations because our data were collected in the summer and others have found that acute anoxic conditions occur in the autumn following algae senescence (Suplee, 2012).

Another way to examine the DO criteria is to look at criteria established to protect against chronic effects from low DO. We did this by comparing the

percentage of times DO fell below the 30-day average minimum DO standard assigned to each site. We acknowledge that these short-term observations (48-72 hour) are not representative of 30-day averages, so some sites that are fully attaining may occasionally fall below these criteria without harming aquatic life. UDWQ currently uses the 30-day average for assessment purposes because it is assumed that this value is more reflective of long-term conditions, so under following these conditions high rates of GPP and CR would correlate strongly with impaired waters. At a minimum, these data suggest that sites with atypically high rates of summertime GPP and CR warrant more intensive investigations into deleterious effects to stream biota.

Our values of GPP (6.0 and 10.0 $gO_2/m^2/day$) and CR (5.0 and 9.0 $gO_2/m^2/day$) A are similar to the suggested rates of GPP and CR proposed by Young et al. (2008) as indicators of river health in New Zealand rivers (7.0 and 9.5 $gO_2/m^2/day$ GPP and CR, respectively). Young and others gathered data from numerous published studies between 1990 and 2006 and then derived thresholds using observations obtained from reference sites to estimate unaltered conditions. Our stressor-response approach along with Young's reference condition approach are part of the growing literature showing multiple lines of evidence that stream metabolism measures are useful measures of stream condition.

B **Physical Covariates.** We found that nutrients were unrelated to metabolic rates at sites where turbidity was greater than 75 ntu, which likely stems from a lack of light reaching the stream benthos. We suggest that stream metabolism is not an appropriate functional indicator to be used at these sites. Nevertheless, at these highly-turbid sites then mean TN (2.41 mg/l) and TP (0.36 mg/l) concentrations were much greater than other States have proposed for numeric nutrient criteria. With nutrient concentrations that exceed even the highest numeric standards nationally, it is unlikely that a functional indicator would be needed to designate impairment at these extremely turbid sites. Secondly, sites with greater than 75 NTU comprise less than three percent of the total stream miles in Utah (DWQ unpublished data). With only a small number of sites that exceed the turbidity criterion for ecosystem metabolism estimation other functional or structural indicators of nutrient enrichment could be used to evaluate these waterbodies for nutrient pollution. These data also highlight the importance of understanding the relative influence of multiple stressors when mitigating for aquatic life degradation, because despite the high concentration of nutrients observed at these sites, excess sedimentation may be a more immediate concern.

To apply the results of our random forest model for management applications we used nonparametric deviance reduction with GPP and CR and physical parameters from random forest models as the explanatory variable to develop a threshold. We found a significant threshold where below 1% slope sites had higher rates of GPP and CR than those above the threshold. Although this relationship is

statistically significant slope was also strongly related to both TN (Pearson correlation r=-0.603) and TP (r=-0.617, data not shown). This relationship is not surprising considering that anthropogenic activities that are known sources of nutrients, including agricultural activities and urban discharges, are more likely to be concentrated at lower gradient, where most of the population resides. The physical parameter we measured that seems to be the clearest covariate influencing stream metabolism appears to be channel shading. Channel shading was also only weakly correlated to in stream TN (Pearson correlation r=-0.245) and TP(r=-0.221, data not shown) indicating that shading effects on GPP and CR are likely not to be confused with those of nutrients. At first it appears that shading having an effect on community respiration is counterintuitive. But in autotrophic systems a large majority of respiration is from autotrophs, in these systems shading would suppress GPP and CR.

Our results indicate that the physical parameters channel slope and shading have the greatest influence on GPP and CR are in agreement with a number of stream metabolism studies. It has been shown that high flow events reduce daily rates of GPP and CR mainly by exporting organic matter accumulated in the stream (Uehlinger et al. 2003, Acuna et al. 2004). Streams with high slope likely accumulate less organic matter and behave like lower gradient streams after high flows. Hill and Dimick (2002) tested the hypothesis that seasonal rates of GPP were attributed to irradiance at the stream surface decreasing due to leaf emergence and not simply temperature. They artificially manipulated shading and found that periphyton photosynthesis declined as irradiance decreased. Bott et al. (2006) and Bernot et al (2010) found that photosynthetic radiation (PAR) was a stronger factor influencing than any measurement of watershed land use. These studies indicate that even under heavily modified land uses (agriculture or urban) GPP can remain relatively low if a healthy riparian corridor is maintained. Or conversely, in ecosystems with little natural riparian vegetation GPP would be highly responsive to anthropogenic increases in nutrients.

Using the significant covariates detected through random forest regression analysis we developed a framework to facilitate covariate effects into management decision making processes (Fig 4.4). We suggest using slope and shading to evaluate Fair GPP and CR groups. The Fair groups for daily GPP and CR rates have slightly increased daily rates over the Good GPP and CR sites but are not large enough to always cause deleterious impacts on minimum oxygen concentrations. We can use our covariates to determine the chance that streams are in this Fair group because of natural intrinsic factors or because of increased nutrient concentrations. We suggest for sites in the Fair GPP and CR categories including slope and shade into the decision making process. Sites where slope is <1% *AND* shading is <11% then it is likely that increased rates of GPP and CR are natural. We would classify these sites as low priority sites. Sites where slope is <1% *OR* shading is less than <11% then there is an intermediate chance that increased rates of GPP

and CR are natural. These sites would be classified as moderate priority sites. Finally sites in the Medium group where slope is >1% and shading is >11% likely have increased GPP and CR from increased nutrient concentrations. These sites would be classified as high priority for sites in the Medium GPP and CR groups (Fig 4.4).

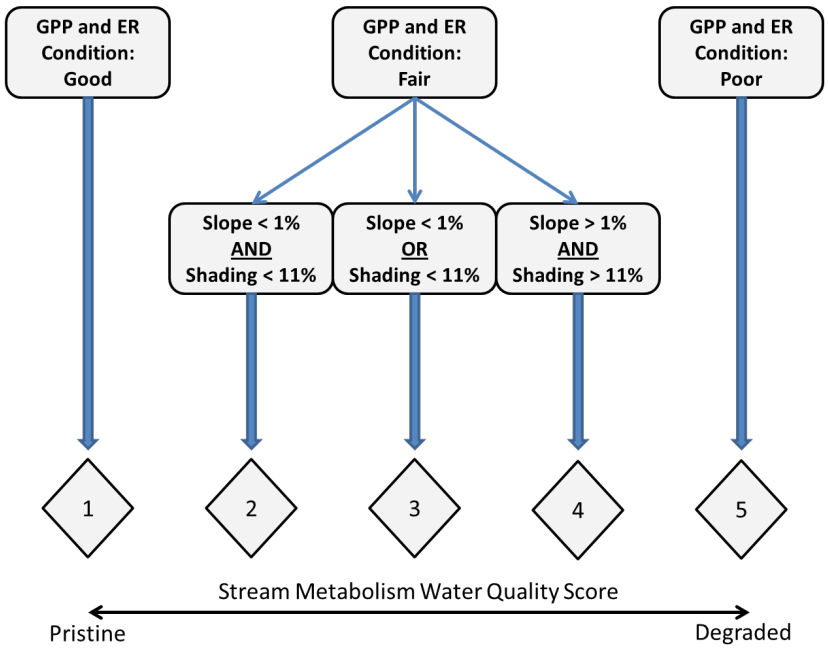

Fig 4.4. Water quality scoring system for evaluating stream metabolism rates (gross primary production and community respiration) in Utah's wadeable streams. We include measurements of physical covariates to help interpret moderately elevated metabolic rates.

Summary and Recommendations

In our study we quantified the relationship between nutrients and stream metabolism and between stream metabolism and in stream minimum dissolved oxygen criteria developed by the Utah Division of Water Quality. We were able to develop thresholds of nutrient concentrations that can be used to screen waterbodies for the potential of having increased rates of GPP and CR. Once these nutrient thresholds are exceeded the Utah DWQ can return to a waterbody and deploy water quality probes and calculate stream metabolism to assess the impacts of nutrient pollution on the stream using the thresholds we developed. Because CR

is the process that produces low oxygen conditions dangerous to biota were are able to connect nutrient concentrations to aquatic life through the processes of stream metabolism.

Literature Cited

Acuna, V., Giorgi, A., Munoz, I., Uehlinger, U. and Sabater, S. 2004. Flow extremes and benthic organic matter shape the metabolism of a headwater Mediterranean stream. *Freshwater Biology.* 49(7):960-971.

Bernot, M.j., Sobota, D.J., Hall, R.O., Mulholland, P.J., Dodds, W.K., Webster, J.R., Tank, J.L., Ashkenas,
L.R., Cooper, L.W., Dahm, C.N., Gregory, S.V., Grimm, N.B., Hamilton, S.K., Johnson, S.L.,
McDowell, W.H., Meyer, J.L., Peterson, B., Poole, G.C., Valett, H.M., Arango, C., Beaulieu, J.J.,
Burgin, A.J., Crenshaw, C., Helton, A.M., Johnson, L., Merriam, J., Niederlehner, B.R., O'Brien,
J.M., Potter, J.D., Sheibley, R.W., Thomas, S.M. and Wilson, A.K. 2010. Inter-regional comparison of land-use effects on stream metabolism. *Freshwater Biology.* 55: 1874-1890.

Bott, T.L., Montgomery D.S., Newbold, J.D., Arscott, D.B., Dow, C.L., Aufdenkampe, A.K., Jackson, J.K. and Kaplan, L.A. 2006. Ecosystem metabolism in streams of the Catskill mountains (Delaware and Hudson River watersheds) and Lower Hudson Valley. *Journal of the North American Benthological Society.* 25(4):1018-1044.

Breiman, L. 2001. Random Forests. *Machine Learning.* 45(1):5-32.

Grace, M.G and Imberger, S. 2006. Stream metabolism: performing and interpreting measurements.
Water Studies Centre Monash University, Murray Darling Basin Commission and New South
Wales Department of Environment and Climate Change. 204pp

Hall, R.O. 2011. Stream metabolism workshop. Utah State University, Logan Utah. Unpublished data.

Hill, B.H., Hall, R.K., Husby, P., Herlihy, A.T. and Dunne, M. 2000. Interregional comparisons of sediment
Microbial respiration in streams. *Freshwater Biology.* 44(3): 213-222.

Houser, J.N., Mulholland, P.J. and Maloney. Catchment disturbance and stream metabolism: patterns in
Ecosystem respiration and gross primary production along a gradient of upland soil and
Vegetation disturbance. Journal of the North American Benthological Society. 24(3):538-552.

Izagirre, O., Agirre, U. Bermejo, M., Pozoz, J. and Elosegi, A. 2008. Environmental controls of whole- stream metabolism identified from continuous monitoring of Basque streams. *Journal of the North American Benthological Society.* 27(2):252.-268

McTammany, M.E., Benfield, E.F. and Webster, J.R. 2007. Recovery of stream ecosystem metabolism

from historical agriculture. Journal of the North American Benthological Society. 26(3):532-545.
Mulholland, P.J., Houser, J.N. and Maloney, K.O. 2005. Stream diurnal dissolved oxygen profiles as
indicators of in-stream metabolism and disturbance effects: Fort Benning as a case study.
Ecological Indicators. 5:243-252.
Odum, H.T. 1956. Primary production in flowing waters. *Limnology and Oceanography.*
1:103-117.
Qian, S.S, King, R.S. and Richardson, C.J. 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modelling. 166:87-97.
R Development Core Team. 2012. R: a language and environment for statistical computing. Vienna,
Austria. www/R-project.org
Sabater, S., Gregory, S.V. and Sedell, J.R. 2002. Community dynamics and metabolism of benthic algae
Colonizing wood and rock substrata in a forested stream. *Journal of Phycology.* 34(4):561-567.
Uehling, U. Kawecka, B. and Robinson, C.T. 2003. Effects of experimental floods on periphyton and stream metabolism below a high dam in the Swiss Alps (River Spol). *Aquatic Sciences.* 65(3):199-209
USEPA. 2009. National Rivers and Streams Assessment: Field Operations Manual.
EPA-841-B-07-009. U.S. Environmental Protection Agency, Washington, DC.
Valderamma, J.C. 1981. The simultaneous analysis of total nitrogen and total phosphorus in natural waters. *Marine Chemistry.* 10(2):109-122.
Van de Bogert, M.C., Carpenter, S.R., Cole. J.J. and Pace, M.L. 2007. Assessing pelagic and benthic metabolism using free water measurements. *Limnology and Oceanography:Methods.* 5:145-155. Young, R.G., Matthaei, C.D. and Townsend, C.R. 2008. Organic matter breakdown and ecosystem metablism: functional indicators for assessing river ecosystem health. *Journal of the North American Benthological Society.* 27(3):605-625.
Young, R.G., and Huryn, A.D. 1999. Effects of land use on stream metabolism and organic matter turnover. *Ecological Applications.* 9:1359-1376.

Appendix A

Appendix A. List of 20 variables used in initial evaluation of physical covariates with random forest regression models. Importance of variables was evaluated using the % increase Mean Squared Error (MSE). Higher MSE indicates that when values in a variable were randomized the model performance declined. Data were obtained from the Utah State University Aquatic Biogeochemistry Laboratory (USU ABL), Utah Unified Public Health laboratories (UPHL), U.S. Geological Survey Stream Stats program (USGS) or the Utah Division of Water Quality Comprehensive Assessment of Stream Ecosystems program (UDWQ).

	Units	GPP %Increase MSE	ER %Increase MSE	Source
Total Nitrogen	mg/l	72.6	69.0	USU ABL
Total Phosphorus	mg/l	73.9	52.1	USU ABL
Turbidity	NTU	41.7	37.0	UPHL
Total Suspended Solids	mg/l	50.0	35.2	UPHL
Channel Shading	%	104.3	69.3	UDWQ
Slope	%	103.1	63.4	USGS
Basin Area	mi^2	58.1	30.4	USGS
Herbaceous Upland	%	69.1	25.7	USGS
Forested Watershed	%	67.5	32.1	USGS
Basin Slope	%	75.3	42.3	USGS
Mean Water Depth	cm	52.6	52.6	UDWQ
Mean Thalweg Depth	cm	32.7	28.0	UDWQ
Bankfull Hieght	cm	69.4	47.3	UDWQ
Channel Incised Hieght	cm	58.6	13.7	UDWQ
Channel Width:Depth Ratio		19.2	30.5	UDWQ
Fine Substrate (<2mm)	%	41.2	19.0	UDWQ
Small Sediment (<16mm)	%	64.5	41.9	UDWQ
Median Particle Size	cm	71.6	46.9	UDWQ
Riffle and Rapid Channel Units	%	38.2	32.5	UDWQ
Riparian Corridor Bare Ground	%	28.0	25.7	UDWQ